Stochastic process (MARKOV CHAINS)

GASMI Abdelkader

April 27, 2025

MARKOV CHAINS

We now consider a sequence of trials whose outcomes, say, $X_1, X_2, \ldots,$ satisfy the following two properties:

1. Each outcome belongs to a finite set of outcomes $S = \{s_1, s_2, \ldots, s_m\}$ called the *state space* of the system. If the outcome on the nth trial is s_i, then we say that the system is in state s_i at time n or at the nth step.
2. The outcome of any trial depends at most upon the outcome of the immediately preceding trial and not upon any other previous outcome. With each pair of states (s_i, s_j), there is given the probability p_{ij} that s_j occurs immediately after s_i occurs.

We say that the stochastic process has the Markov property if

$$P(X_{n+1} = s_i / X_n; X_{n-1}; \ldots; X_0) = P(X_{n+1} = s_i / X_n), \qquad (1)$$

for all $n \geq 0$ and $s_i \in S$. We will discuss the case when S is a finite set. In this case, we let S=[m] for some positive integer m.

A stochastic process with the Markov property is called a Markov chain. Such a stochastic process is called a *(finite) Markov chain*. The numbers p_{ij}, called the *transition probabilities*, can be arranged in a matrix:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix},$$

called the *transition matrix*.

Thus, with each state s_i, there corresponds the ith row $(p_{i1}, p_{i2}, \ldots, p_{im})$ of the transition matrix P. If the system is in state a_i, then this row vector represents the probabilities of all the possible outcomes of the next trial, and so it is a *probability vector*.

Theorem (1)

The transition matrix P of a Markov chain is a stochastic matrix.

Example (1)

A man either drives his car or takes a train to work each day. Suppose he never takes the train two days in a row; but if he drives to work, then the next day he is just as likely to drive again as he is to take the train.

The state space of the system is $\{t\,(\text{train}), d\,(\text{drive})\}$. This stochastic process is a Markov chain since the outcome on any day depends only on what happened the preceding day. The transition matrix of the Markov chain is:

$$P = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

The first row of the matrix corresponds to the fact that he never takes the train two days in a row and so he definitely will drive the day after he takes the train. The second row of the matrix corresponds to the fact that the day after he drives he will drive or take the train with equal probability.

Example (2)

Random Walk with Reflecting Barriers: A man is at an integral point on the z-axis between the origin 0 and the point 5. He takes a unit step to the right with probability p or to the left with probability $q = 1 - p$, unless he is at the origin, where he takes a step to the right to 1, or at the point 5, where he takes a step to the left to 4. Let X_n denote his position after n steps.

The transition matrix is:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 \\ 0 & 0 & q & 0 & p & 0 \\ 0 & 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Each row of the matrix, except the first and last, corresponds to the fact that the man moves from state s_i to state s_{i+1} with probability p or back to state s_{i-1} with probability $q = 1 - p$. The first row corresponds to the fact that the man must move from state s_0 to state s_1, and the last row corresponds to the fact that the man must move from state s_5 to state s_4.

Higher Transition Probabilities

The entry p_{ij} in the transition matrix P of a Markov chain is the probability that the system changes from the state s_i to the state s_j in one step: $s_i \to s_j$.

Question: What is the probability, denoted by $p_{ij}^{(n)}$, that the system changes from the state s_i to the state s_j in exactly n steps:

$$s_i \to s_{k_1} \to s_{k_2} \to \cdots \to s_{k_{n-1}} \to s_j?$$

The following theorem answers this question; here the $p_{ij}^{(n)}$ are arranged in a matrix $P^{(n)}$, called the n-step transition matrix:

Theorem (2)

Let P be the transition matrix of a Markov chain process. Then the n-step transition matrix is equal to the nth power of P; that is,

$$P^{(n)} = P^n.$$

Now suppose that, at some arbitrary time, the probability that the system is in state a_i is p_i. We denote these probabilities by the probability vector

$$\mathbf{p} = (p_1, p_2, \ldots, p_m),$$

which is called the *probability distribution* of the system at that time.

In particular, we define:

$$\mathbf{p}^{(0)} = (p_1^{(0)}, p_2^{(0)}, \ldots, p_m^{(0)}),$$

as the *initial probability distribution*, i.e., the distribution when the process begins. Similarly, we denote the nth step probability distribution by:

$$\mathbf{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, \ldots, p_m^{(n)}),$$

i.e., the distribution after the first n steps.

Theorem (3)

Let P be the transition matrix of a Markov chain process. If $\mathbf{p} = (p_i)$ is the probability distribution of the system at some arbitrary time, then:

$$\mathbf{p}P$$

is the probability distribution of the system one step later, and:

$$\mathbf{p}P^n$$

is the probability distribution of the system n steps later.

In particular,

$$\mathbf{p}^{(1)} = \mathbf{p}^{(0)}P,\ \mathbf{p}^{(2)} = \mathbf{p}^{(1)}P,\ \mathbf{p}^{(3)} = \mathbf{p}^{(2)}P, \ldots,\ \mathbf{p}^{(n)} = \mathbf{p}^{(n-1)}P$$

Example (3)

Consider the Markov chain of Example (1), whose transition matrix is:

$$P = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Here, t represents the state of taking a train to work, and d represents the state of driving to work.
Using the properties of Markov chains, the probability that the system changes from state t to state d in exactly 4 steps is:

$$p_{td}^{(4)} = \frac{5}{8}.$$

Similarly:

$$p_{tt}^{(4)} = \frac{3}{8}, \quad p_{dd}^{(4)} = \frac{5}{16}, \quad p_{dt}^{(4)} = \frac{11}{16}.$$

Now suppose that on the first day of work, the man tossed a fair die and drove to work if and only if a 6 appeared. In other words, the initial probability distribution is:

$$\mathbf{p}^{(0)} = \left(\frac{5}{6}, \frac{1}{6} \right).$$

After 4 days, the probability distribution becomes:

$$\mathbf{p}^{(4)} = \mathbf{p}^{(0)} P^4 = \left(\frac{35}{96}, \frac{61}{96} \right).$$

Thus:

$$p_t^{(4)} = \frac{35}{96}, \quad p_d^{(4)} = \frac{61}{96}.$$

Stationary Distribution of Regular Markov Chains

Suppose that a Markov chain is regular, i.e., that its transition matrix P is regular. By Theorem (fixed points), the sequence of n-step transition matrices P^n approaches the matrix T, whose rows are each the unique fixed probability vector $\mathbf{t}$ of P. Hence, the probability $p_{ij}^{(n)}$ that state s_j occurs for sufficiently large n is independent of the original state s_i, and it approaches the component t_j of $\mathbf{t}$. In other words:

Theorem (4)

Let the transition matrix P of a Markov chain be regular. Then, in the long run, the probability that any state a_j occurs is approximately equal to the component t_j of the unique fixed probability vector $\mathbf{t}$ of P.

Thus, we see that the effect of the initial state or the initial probability distribution of the process wears off as the number of steps of the process increases. Furthermore, every sequence of probability distributions approaches the fixed probability vector $\mathbf{t}$ of P, called the *stationary distribution* of the Markov chain.

Example (4)

Consider the Markov chain process of Example 7.12, whose transition matrix is:

$$P = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

To find the stationary distribution $\mathbf{t} = (t_t, t_d)$ of this Markov chain, we solve the equations:

$$\mathbf{t}P = \mathbf{t}, \quad \text{and} \quad t_t + t_d = 1.$$

Expanding the equation $\mathbf{t}P = \mathbf{t}$:

$$\begin{bmatrix} t_t & t_d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} t_t & t_d \end{bmatrix}.$$

This simplifies to the system of linear equations:

$$\frac{1}{2} t_d = t_t, \quad t_t + \frac{1}{2} t_d = t_d.$$

Solving these equations by using $t_d + t_t = 1$, we find:

$$t_t = \frac{1}{3} \quad and \quad t_d = \frac{2}{3}.$$

Thus, the stationary distribution of the Markov chain is:

$$\mathbf{t} = \left(\frac{1}{3}, \frac{2}{3} \right).$$

Absorbing States

Definition

A state s_i of a Markov chain is called *absorbing* if the system remains in the state s_i once it enters there. Thus, a state s_i is absorbing if and only if the ith row of the transition matrix P has a 1 on the main diagonal and zeros everywhere else. (The main diagonal of an $n \times n$ matrix $A = (_{ij})$ consists of the entries $a_{11}, a_{22}, \ldots, a_{nn}$.)

Example (5)

Suppose the following matrix is the transition matrix of a Markov chain:

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The states a_2 and a_5 are each absorbing since the second and fifth rows have a 1 on the main diagonal and zeros everywhere else.